Rudolph, Stephan

Working Paper

The gravity equation with micro-founded trade costs

Dresden discussion paper series in economics, No. 11/09

Provided in cooperation with:
Technische Universität Dresden

TU Dresden
Faculty of Business and Economics

Dresden Discussion Paper Series in Economics

The Gravity Equation with Micro-founded Trade Costs

STEPHAN RUDOLPH

Dresden Discussion Paper in Economics No. 11/09

ISSN 0945-4829

Address of the author(s):

Stephan Rudolph
Technische Universität Dresden
Helmholtzstrasse 10
01069 Dresden
Germany

e-mail : Stephan.Rudolph@tu-dresden.de

Editors:

Faculty of Business and Economics, Department of Economics

Internet:

An electronic version of the paper may be downloaded from the homepage:
http://rcswww.urz.tu-dresden.de/wpeconomics/index.htm

English papers are also available from the SSRN website:
http://www.ssrn.com

Working paper coordinator:

Dominik Maltritz
e-mail: wpeconomics@mailbox.tu-dresden.de

Dresden Discussion Paper in Economics No. 11/09

The Gravity Equation with Micro-founded Trade Costs

Stephan Rudolph
Technische Universität Dresden
01069 Dresden
Stephan.Rudolph@tu-dresden.de

Abstract:

Gravity Equations are broadly used to estimate the impacts of trade impediments on trade flows. It is often stated that results are implausibly high. In theoretical foundations of the gravity equation, trade costs usually enter as "iceberg-melting-costs". This paper offers an alternative approach to model trade costs. From a microeconomic point of view, trade costs should depend on trade input prices and -- which is new -- the underlying trade volume. If trade costs are determined by the trade volume, and average trade costs are falling with the trade volume (e.g. due to economies of scale in the trade sector), empirical results from gravity equations are likely to be biased.

JEL-Classification: F10, C51

Keywords: gravity equation, trade costs, estimation bias

Acknowledgements:
I am grateful to Amit Biswas, Udo Broll, Georg Hirte, Dennis Novy and Marco Wagner for useful comments.

1 Introduction

The gravity equation is a widely used tool in international economics to explain and estimate trade flows. It is also broadly used to estimate the effects of policy variables like the membership in a trade- or monetary union on bilateral trade. The concept behind the gravity equation is to regress bilateral trade flows on the exporter's and importer's economic sizes and trade barriers between them in a log-linear form. The name comes from the analogy to Issac Newton's law of gravity, where the force of attraction between two bodies depends on the bodies' masses divided by the squared distance. The idea to take a similar form to explain trade volumes between to countries was developed by Tinbergen (1962) and Pöyhönen (1963), independently of each other.

After rising criticism that the gravity equation is a purely intuitive and not theoretically founded empirical tool, Anderson (1979) was one of the first who developed a theoretical framework to derive the gravity equation. He uses an expenditure system where countries are exogenously endowed with a certain GDP, so that GDP is not the result of an underlying production function. Anderson and van Wincoop (2004) later call this approach "separable trade theory", because trade is separated from determinants like technology or factor endowment. The work of Bergstrand recognizes these determinants by using explicit production frontiers with constant elasticities of transformation to derive a theoretically founded gravity equation (see Bergstrand, 1985, 1989, 1990). Another foundation based on the theory of market structure and foreign trade was developed by Helpman and Krugman (1985, chapter 8).

These works build up the gravity equation on the so called new-trade-theory introduced by Krugman (1979), where increasing returns and imperfect competition are the reasons for trade. Models by Deardorff (1998) and Eaton and Kortum (2002) showed that also the traditional classical/neo-classical trade models (Ricardo and Heckscher-Ohlin), which point out comparative advantages as the reason for trade, can be used as a theoretical basis of the gravity equation. Recently, the so called new-new-trade-theory developed by Melitz (2003) attracts a lot of interest. This theory argues that first of all firm characteristics and not country characteristics lead to trade. Helpman, Melitz, and Rubinstein (2008) introduced this approach into the gravity literature.[1]

One important difference between the gravity equation and nearly all theoretical models of international trade is that the gravity equation allows for a measure of trade costs or trade barriers. In standard international trade theory, trade costs are not considered. The theoretical gravity equation explains bilateral trade by economic sizes and trade costs. Trade costs usually enter this equation as "iceberg-melting-costs" introduced by Samuelson (1954): If a certain good is sent from one country to another, this good looses a *fixed* part of its value. Iceberg-costs can be interpreted as an ad valorem tariff equivalent for trade costs. This way of modeling trade costs is very common in theoretical models, because it is quite tractable.

Nevertheless, Grossman (1998) criticizes the iceberg-approach in theoreti-

[1] Helpman, Melitz, and Rubinstein (2008) furthermore explain missing trade flows between country pairs by technological differences between firms. They argue that only firms with high levels of productivity will be able to compete on international markets and so become exporters.

cal gravity models as a "technology for shipping tomatoes" and raised the suspicion that a wrong consideration of trade costs in gravity frameworks could be a reason for what Obstfeld and Rogoff (2001) later posed as one of their six puzzles of international macroeconomics: the problem that the estimated coefficients of border and distance effects on trade have unexpected high values.[2] But recent theoretical contributions that help to improve the interpretation of the empirical outcomes of gravity equations, like Anderson and van Wincoop (2003), also use the concept of iceberg-costs to insert trade costs into their model.[3] It is worth to note that the very recent studies by Helpman, Melitz, and Rubinstein (2008) or Chaney (2008) use an augmented variation of iceberg-costs with an additional fixed mark up for shipping one unit from one certain country to another.[4] But in this augmentation, (average) trade costs become only a function of the underlying factory price and not of the trade volume at all. In this paper I argue, using a simple micro-economic model, that (average) trade costs should be determined by the trade volume (price multiplied by quantity) and not be constant (like iceberg-costs imply) or only depend on the underlying factory prices (like in recent studies).

[2] The probably most cited example is McCallum (1995), who estimated that the border between Canada and the United States makes trade between a certain Canadian province and another Candian province by a factor 22 (2,200 percent) higher than trade between this Canadian province and an U.S.-state of the same economic size and the same distance.

[3] The innovation by Anderson and van Wincoop (2003) is that trade barriers between two countries must be seen relatively to the trade barriers with all other barriers of these two countries. Their approach is the basis for this analysis.

[4] As a result, Chaney (2008) yields an endogeneous elasticity of the trade volume with respect to trade barriers.

The aim of this paper is to bring trade costs adequately into a theory based gravity equation. Because iceberg-costs can be interpreted as fixed average costs, they are independent from the underlying trade volume. Since there are economies of scale in trade this assumption is inadequate: the higher the trade volume between two countries, the lower should be the cost of sending one (composite) unit of the export volume from the one to the other country, since economies of scale cause declining average costs. This suggestion results from microeconomic theory. It leads to an endogeneity problem in empirical gravity equations and hence to a bias in the estimated parameters. Under certain circumstances, this bias can be a contribution to explain implausibly high estimates for border effects in gravity frameworks.

The paper is structured as follows. Section 2 introduces the theoretical derivation of the gravity equation by Anderson and van Wincoop (2003) with trade costs modeled as iceberg-costs. Section 3 offers an alternative approach to model trade costs. If trade volume is seen as the output of a trade sector, microeconomic theory reveals that the according trade costs depend on input prices but also on the trade volume. Economies of scale in this trade sector, which is presumable according to several empirical studies, lead to decreasing average trade costs in trade volume. Section 4 brings these micro-founded trade costs into the theory based gravity equation and extracts the bias term that might influence empirical studies using the gravity equations. Section 5 concludes.

2 The Theory Based Gravity Equation

This section recalls the theory based derivation of the gravity equation introduced by Anderson and van Wincoop (2003). Consider a world with many countries. The GDP of these countries is given exogenously. Each country i's total production Y_i can be seen as a specific good of this country – the so called Armington assumption (Armington, 1969). The intuition of this assumption is that consumers – to give an example – don't care if it is a car or an apple, but they care where the commodity has been produced.[5] Consumers over the world are supposed to have the same preferences. An exporting country will be denoted with i, an importing country with j.

Following Anderson and van Wincoop (2003) trade costs enter the model as iceberg-costs. Iceberg-costs are a fixed exogenously given markup ("iceberg-factor") t_{ij} onto the factory price p_i, so that the price of the (composite) commodity of country i payed in country j is $p_{ij} = t_{ij} \cdot p_i$. The price of the commodity from i is in country j higher by the factor t_{ij} than in the country of origin i due to trade costs. It is assumed that $t_{ij} > 1$ for all countries $j \neq i$ and that the domestic trade cost factor $t_{ii} = 1$. This is to ensure that commodities are more expensive abroad than on the domestic market. Modeling trade costs in this way leads to three properties. First, since the trade volume including transport costs (gross trade volume) is $T_{ij} = t_{ij} \cdot p_i \cdot c_{ij}$ with quantity c_{ij} sent from i to j, the trade volume can be decomposed into total trade costs $(t_{ij} - 1) \cdot p_i \cdot c_{ij}$ plus transport cost exclusive (net) trade

[5] This assumption is used for simplicity. Anderson and van Wincoop (2004) show the same results with many goods per country. Also see Deardorff (1998) for a discussion of the case of many goods and relaxing the Armington assumption.

volume $p_i \cdot c_{ij}$.[6] Second, it can be shown that a fraction $(t_{ij} - 1)/t_{ij}$ of the amount of goods shipped from i to j is lost in transport.[7] Finally, iceberg-costs are a measure for average trade costs and not for total trade costs. This property is important for the message of this paper. Obviously, the iceberg-factor can be denoted as gross trade volume devided by net trade volume:

$$t_{ij} = \frac{p_{ij} \cdot c_{ij}}{p_i \cdot c_{ij}}.$$

This implies that t_{ij} is nothing else than the tariff-equivalent factor for bringing a value of \$ 1.00 of country i's composite export good to country j. Therefore, iceberg-cost-factor t_{ij} is nothing else than a *per-unit-cost* of trade.

Consider an importing country j. Recall that consumers over the world have identical preferences, so that preferences of the consumers in country j can be represented by the CES-utility-function

$$U_j = \left(\sum_i \beta_i \cdot c_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{1}$$

Here, c_{ij} is the quantity of i's commodity imported by j (including country j's domestic consumption c_{jj}), β_i is a distribution parameter to weight the preference of the representative consumer for country i's composite good and σ is the elasticity of substitution between all goods of the world. This elasticity of substitution is assumed to be $\sigma > 1$, meaning that there is

---

[6] To bring this mathematically into one line: $T_{ij} = p_{ij} \cdot c_{ij} = t_{ij} \cdot p_i \cdot c_{ij} = (t_{ij} - 1) \cdot p_i \cdot c_{ij} + p_i \cdot c_{ij}$. The last expression shows that T_{ij} equals total trade costs (first summand) plus the net trade volume (second summand).

[7] Assume for simplicity that $p_i = 1$ and $c_{ij} = 1$ and e.g. $t_{ij} = 1.25$. This means, country i must send 1.25 units to j so that one unit arrives. In this case a fraction $0.25/1.25 = 0.2$ of the trade volume sent by country i would be lost.

a substitutive relationship between the single commodities by the different countries.[8] The budget constraint of country j postulates that its GDP Y_j must equal the expenditure of country j on all goods of all countries i (inclusive the good of country j itself, $T_{jj} = p_{jj} \cdot c_{jj}$):

$$Y_j = \sum_i p_{ij} \cdot c_{ij}, \tag{2}$$

with p_{ij} as the price of i's commodity in country j. The factory price of i's commodity, meaning the price without any trade costs, will be denoted with p_i.

Maximizing country j's utility function subject to its budget constraint yields the demand function and multiplying both sides of this demand function by p_{ij} yields the import function

$$T_{ij} = \beta_i^{\sigma} \cdot \left(\frac{t_{ij} \cdot p_i}{P_j} \right)^{1-\sigma} \cdot Y_j, \tag{3}$$

with $T_{ij} = t_{ij} \cdot p_i \cdot c_{ij}$ being the gross value of imports of j from i and

$$P_j = \left(\sum_i \beta_i^{\sigma} t_{ij}^{1-\sigma} p_i^{1-\sigma} \right)^{1/(1-\sigma)} \tag{4}$$

being a CES-price-index of country j.

Now, consider an exporting country i. In a general equilibrium with clear markets, GDP of country i must equal the sum of all exports (including the export into i itself – i's intra-trade T_{ii}). Combining this general equilibrium

[8] In a review of empirical literature, Anderson and van Wincoop (2004) point out that this elasticity of substitution σ lies between 5 and 10.

condition with equation (3) gives:

$$
\begin{aligned}
Y_i &= \sum_j T_{ij} \\
&= \sum_j \beta_i^\sigma \cdot \left(\frac{t_{ij} \cdot p_i}{P_j}\right)^{1-\sigma} \cdot Y_j \\
&= \beta_i^\sigma p_i^{1-\sigma} \cdot \sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \cdot Y_j \\
&= \beta_i^\sigma p_i^{1-\sigma} \cdot Y_w \sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \cdot s_j \\
&= \beta_i^\sigma p_i^{1-\sigma} \cdot Y_w \cdot \Pi_i^{1-\sigma},
\end{aligned} \tag{5}
$$

with $Y_w = \sum_j Y_j$ being the world's GDP, $s_j = Y_j / Y_w$ being country j's share of world GDP and

$$
\Pi_i \equiv \left(\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \cdot s_j\right)^{1/(1-\sigma)} \tag{6}
$$

being a measure for country i's "multilateral resistance", as it is an index for trade costs of country i with all countries (summed over j).

Solving equation (5) for the scaled prices ($\beta_i^\sigma p_i^{1-\sigma}$) and using this for the CES-index (4) yields the multilateral resistence term for country j:

$$
P_j = \left(\sum_i \left(\frac{t_{ij}}{\Pi_i}\right)^{1-\sigma} \cdot s_i\right)^{1/(1-\sigma)}. \tag{7}
$$

Substituting the solution of equation (5) for the scaled prices ($\beta_i^\sigma p_i^{1-\sigma}$) into the import volume function (3) finally gives the theory based gravity equation

$$
T_{ij} = \frac{Y_i \cdot Y_j}{Y_w} \cdot \left(\frac{t_{ij}}{\Pi_i \cdot P_j}\right)^{1-\sigma}. \tag{8}
$$

Note, that (8) includes trade costs on both sides. For later purposes of this paper, it will be useful to consider trade flows without trade costs. The according gravity equation without trade costs follows from dividing (8) by t_{ij}:

$$T_{ij}^0 = \frac{Y_i \cdot Y_j}{Y_w} \cdot t^{-\sigma} \cdot (\Pi_i \cdot P_j)^{\sigma-1}\,, \tag{9}$$

where T_{ij}^0 denotes trade cost adjusted trade flows (T_{ij}/t_{ij}) or the *net* trade volume, T_{ij} the *gross* trade volume.[9]

As long as the elasticity of substitution between the countries' goods, σ, is larger than 1, higher bilateral iceberg-trade-costs lower the bilateral trade volume. Since factor t_{ij} can be interpreted as the cost of bringing a value of \$ 1.00 from country i to j, a kind of average trade costs, it follows from gravity equation (8) and (9): the higher the average trade costs between two countries, the lower the trade volume. Considering factor t_{ij} not as some undefined measure of trade costs but explicitly as the average trade cost value will be a central message of this paper.

The second contribution of the Anderson/van-Wincoop-model is that these average trade costs do not purely enter the gravity equation (like in older versions), but they must be seen relatively to the product of the multilateral resistances of the trading partners: It is not enough to consider average trade costs between two countries, bilateral average trade costs *relative* to all other trading partners must enter the model. Several studies show that controlling for these multilateral resistances lowers implausibly high border effects (see Hummels, 1999; Rose and van Wincoop, 2001; Anderson and van Wincoop, 2003, and others).

[9] If trade costs were only costs of insurance and trade, T_{ij} could be called c.i.f.-trade-volume and T_{ij}^0 f.o.b.-trade-volume, but in the context of this paper, trade costs can play a much broader role.

In empirical applications, equation (9) can be estimated as a log-linear function:

$$\ln T^0_{ij} = \alpha_0 + \alpha_1 \ln Y_i Y_j + \alpha_2 \ln t_{ij} + FE_i + FE_j + \tilde{u}_{ij}, \tag{10}$$

where $\tilde{u}_{ij}$ is the log of the error term and FE_i, FE_j are fixed effects to capture i's and j's multilateral resistances, respectively. As theory from equation (9) implies, $1/Y_w$ is captured by intercept α_0, α_1 should be approximately 1 and α_2 should represent the negative value of the elasticity of substitution, $-\sigma$. While data for T^0_{ij} and Y_i, Y_j is available, there is no satisfying data source for average trade costs and the (even unobservable) multilateral resistance terms. Average trade costs are usually proxied by distance and further variables that control for proximity of two trading countries. Frequently used variables for these issues are geographical variables as country area, coast length, exchange rate volatility as well as dummies for common border, common language and so on. The problem with the multilateral resistance terms is usually solved by using country fixed effects: importer and exporter dummies.[10] This technique tends to reduce implausibly high parameters for diverse border effects.

3 A Micro-founded Form of Trade Costs

In the setup with iceberg-costs of the previous section, t_{ij} is a constant factor that represents average costs of trade. This factor is not directly measurable and usually is proxied by distance and several control variables, for example dummies for common border and common language. But, since t_{ij} denotes

[10] Anderson and van Wincoop (2003) also compute the resistance term numerically. It appears that results of this numerical approach are similar to the use of country fixed effects.

average costs for trade, it has never been modeled to my knowledge as a micro-founded average cost function. From microeconomic theory, it is well known that an average cost function not only depends on cost factors like factor prices but also on the produced quantity. And therefore, I argue that average trade costs are dependent on the trade volume.

If a bilateral net trade volume is the output of a trade sector production function, we can denote it as a function of input vector $x_{ij}^k = (x_{ij}^1, \dots, x_{ij}^K)$:

$$T_{ij}^0 = T_{ij}^0(x_{ij}^k). \tag{11}$$

An input x_{ij}^k in this context can mean, for example, shipping one good via ocean or air, paying for tariffs, translating contracts, and so on. Now, let $w_{ij}^k = (w_{ij}^1, \dots, w_{ij}^K)$ be the vector of input prices. Minimizing trade costs $\sum_k w_{ij}^k \cdot x_{ij}^k$ subject to a given net trade volume, presumed that second order conditions hold, yields the trade cost function

$$TC_{ij} = TC_{ij}(w_{ij}^k, T_{ij}^0). \tag{12}$$

Dividing both sides by T_{ij}^0 yields average trade costs,

$$\overline{TC}_{ij} = \frac{TC_{ij}}{T_{ij}^0} = \overline{TC}_{ij}(w_{ij}^k, T_{ij}^0) = t_{ij}(w_{ij}^k, T_{ij}^0). \tag{13}$$

These average trade costs $\overline{TC}_{ij}$ describe the costs of bringing a value of \$ 1.00 from country i's composite trade volume to country j. This is equal to the interpretation of the iceberg-factor t_{ij}. Thus, the iceberg-factor becomes endogenous. As long as there are economies of scale in the trading sector, e.g. caused by the presence of fixed costs of infrastructure, average cost function (13) will decline with the bilateral trade volume: The more two countries trade with each other, the lower are the average bilateral trade costs. The result is the presumption that $\partial t_{ij}/\partial T_{ij}^0 < 0$.

Why should there be economies of scale in the trade sector? To answer this question we first need to know what trade costs are. According to this paper, trade costs are the costs for bringing a product from a home market on a foreign market. Following Anderson and van Wincoop (2004) they can be subdivided into three different kinds of trade costs: (a) transport costs, (b) border-related trade barriers and (c) costs for retail and wholesale on the foreign market.

Transport costs are the costs for shipping goods. They can be separated into direct transport costs, the so called costs of insurance and freight (c.i.f.), and indirect transport costs, which include holding costs for goods in transit, inventory costs due to buffering the variability of delivering dates, preparation costs associated with shipment size and other costs. Hummels (2007) argues that the most important technologies of shipping goods between countries are ocean and air shipping. As one reason for this he points out that only one quarter of the world's trade volume takes place between countries that share a common border. There are several approaches to capture trade costs with empirical data (see Hummels, 1999; Limao and Venables, 2001; Redding and Venables, 2002; Hummels, 2007, for example), although especially indirect transport costs are hardly observable. In gravity equations, transport costs are usually proxied by distance between the capitals or economic centers of two trading countries.

Border-related trade barriers are trade impediments which occur between countries due to political, currency, language, cultural and other reasons. The problem with these barriers is that there are many unobservable and probably even unknown effects. Some barriers are observable, like e.g. tariffs, currency volatilities and so on, but there are data limitations to the political barriers, as Anderson and van Wincoop (2004, section 2.1.1) criticize. Notably, it is very service-intensive to overcome barriers like language, mentality, bureaucracy and so on. In gravity equations, border-related trade effects are usually controlled by a set of dummy variables for common properties of the countries.

Costs for wholesale and retail have to be borne by both suppliers from the country and suppliers from abroad, which export to this country. Since these costs differ between countries, they are likely to enter the exporter's decisions. Wholesale and retail costs a captured in gravity equations via price indices (following Baier and Bergstrand, 2001, and the earlier work by Bergstrand) or, more commonly, multilateral resistance terms and/or therefore country fixed effects (Anderson and van Wincoop, 2003).

In summary, per unit costs of bringing goods from one country, t_{ij}, into another country should depend on (a) transport and (b) border effect cost, while (c) costs for wholesale and retail should be captured by individual country effects (Π_i, P_j).

The transport sector typically uses fix cost intensive infrastructure: harbors, airports, rail track networks, road systems. Limao and Venables (2001) find that infrastructure plays an important role for the determination of transport costs, especially for landlocked countries. As market power indicates a presence of fixed costs and economies of scale, e.g. Hummels, Lugovskyy, and Skiba (2009) find evidence for market power and price discrimination in the ocean shipping industry.

Furthermore, work by Hummels (and co-authors) indicates that the usage of fixed iceberg-melting-costs is an inappropriate measure for transport costs.

Hummels and Skiba (2002, pp. 2–6) give a detailed discussion of the sources of scale economies in the transport sector. As an introductory example, they argue that shipping goods from Ivory Coast to the U.S. East Coast is twice as high as shipping goods from Japan to the U.S. West Coast, although distance is the same in both cases.

On the other side, costs for border-related effects are also likely to have economies of scale. As noted above, overcoming border related effects can be closely related to services. Here, social networks, communication networks and many more factors should play an important role (see Jones and Kierzkowski, 1990, for a discussion of the particular case in which trade goods are produced in a fragmented industry) and there should be a relationship between costs for these service networks and the underlying trade volume, similar to a technology with fix costs. If there is more trade, per unit costs for translations, filling out forms, overcoming bureaucracy and others should be lower.

All of these arguments lead to the proposition that average trade costs should depend on the bilateral trade volume and that the relationship between them is negative.

4 Endogenous Trade Costs in Gravity Equations

If bilateral trade costs depend on the underlying trade volume, an endogeneity problem may bias estimations from gravity equations. After inserting the endogenous average trade costs into the gravity equation, we can extract a functional term that shows the bias and discuss it.

The first step is to bring endogenous trade costs into a functional form that is suitable to empirical studies using the gravity equation. Grossman (1998) suggests a log-linear form for (per unit) trade costs that respects distance D_{ij} and a vector of border effects b_{ij}:

$$\ln t_{ij} = \beta_0 + \beta_1 b_{ij} + \beta_2 \ln D_{ij}. \tag{14}$$

If t_{ij} is also a function of the trade volume, T^0_{ij} could enter this log-linear form:

$$\ln t_{ij} = \beta_0 + \beta_1 b_{ij} + \beta_2 \ln D_{ij} + \beta_3 \ln T^0_{ij}. \tag{15}$$

The second step is to insert this trade cost modeling into the empirical gravity equation. Substituting equation (15) into equation (10) yields:

$$\begin{aligned} \ln T^0_{ij} = \ \alpha_0 \ & + \alpha_1 \ln Y_i Y_j + \alpha_2(\beta_0 + \beta_1 b_{ij} + \beta_2 \ln D_{ij} + \beta_3 \ln T^0_{ij}) \\ & + FE_i + FE_j + \tilde{u}_{ij}. \end{aligned} \tag{16}$$

In equation (16), net trade volume T^0_{ij} appears on both sides and since it has got an impact on trade costs ($\beta_3 \neq 0$) it should cause a bias.

The third step is to extract the bias term. Equation (16) can easily be solved for $\ln T^0_{ij}$:

$$\ln T^0_{ij} = \frac{1}{1-\alpha_2\beta_3} \cdot \left[\begin{array}{ll} \alpha_0 & +\alpha_1 \ln Y_i Y_j + \alpha_2(\beta_0 + \beta_1 b_{ij} + \beta_2 \ln D_{ij}) \\ & +FE_i + FE_j + \tilde{u}_{ij} \end{array} \right]. \quad (17)$$

The bias of ignoring endogeneity of trade costs thus is given by the fraction $1/(1-\alpha_2\beta_3)$. From a theoretical point of view, α_2 in equation (10) can be interpreted as $-\sigma$, as can be seen from equation (9). As σ is assumed to be larger than 1, α_2 will be negative. If there are economies of scale in the trade sector and per unit trade costs decrease in trade volume, β_3 should also be negative. With negative α_2 and β_3, $\alpha_2\beta_3$ must be positive. If $\alpha_2\beta_3$ lies between 0 and 1, the bias is positive and larger than one. This would imply that coefficients are overestimated as long as trade costs are not considered to be endogenous. With $\alpha_2\beta_3$ converging against 1, the bias grows exponentially against infinity. At $\alpha_2\beta_3 = 1$ there is no solution for the bias. If $\alpha_2\beta_3$ is larger than 1, the fraction becomes negative. This would imply that the signs of the estimated effects are changed – which would lead to implausible estimates and that would be contradictory to the success of the gravity equation. The bias term is plotted in appendix A.1.

According to Anderson and van Wincoop (2004), most studies of substitutability of internationally traded goods estimate substitution elasticities σ between 5 and 10, so that α_2 should lie in the interval $[-5; -10]$. If β_3 is exactly $-1/5$ or $-1/10$, respectively, the bias would be indefinite. If the absolute value of β_3 is higher than these values, the bias would convert the parameters' signs and the gravity equation would not be as famous as it is. If the absolute value of β_3 is lower than these values, but greater than 0, estimated parameters are biased upwards. If $\beta_3 = 0$, which is implicitly assumed in gravity works until now, the fraction would be one, implying that there is no bias. If $\beta_3 > 0$, trade volume would have a positive effect on average trade costs which is hard to explain in a sector that is likely to deal with fixed costs. In this case, standard estimations with the gravity equation would be underestimated and the discussion about the border puzzle would go into the wrong direction. The bias term with a given σ is plotted in appendix A.2.

If trade volumes have an impact on per unit trade costs and if the discussion about the puzzle of the implausibly high home bias in gravity equations is on the right path, β_3 must lie between 0 and the inverse value of $-\sigma$. This indicates that average trade costs' elasticity of the trade volume should be low, but not zero.

5 Conclusion

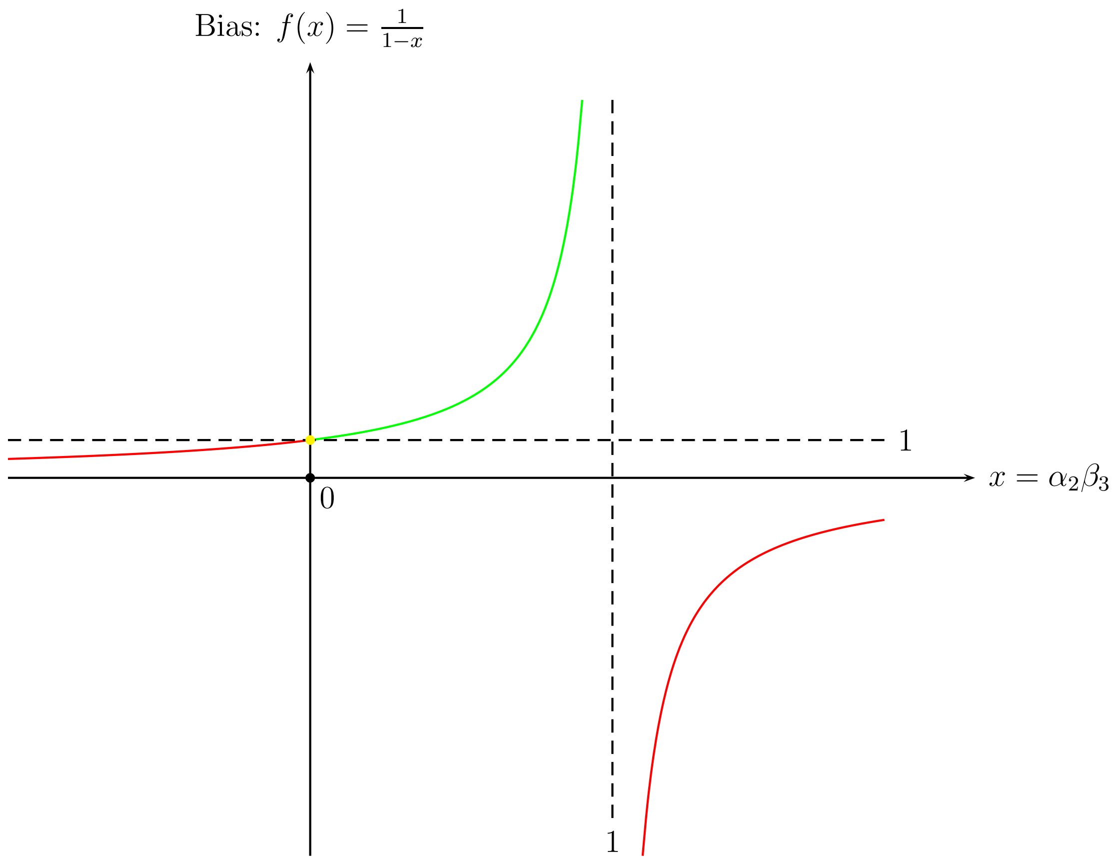
This paper introduces an alternative form of bringing trade costs into a theory based gravity equation. The main argument is that, from a microeconomic point of view, trade costs should not be independent from the underlying trade volume and, since there are economies of scale in the trade sector, average trade costs should decline with the underlying trade volume. If this relationship is not controlled in empirical studies using the gravity equation, the results might be biased. The analysis of the bias shows that the impact of trade volume on average trade costs should be very inelastic, else the results of gravity studies are hard to explain. But if the impact is significant, the bias might be an explanation for overestimates, which could be a contribution to the discussion about the "border-puzzle".

Another interesting outcome of this paper is that input prices of trade factors play an important role in a micro-founded trade cost function. Usually, trade barriers are proxied by distance and some dummy-variables. But micro-founded cost theory postulates that proxies for input prices should enter the trade cost terms to reflect the aggregated technology of the trade sector. Notably, Brun, Carrère, Guillaumont, and de Melo (2005) gain a higher explanatory content with additional price variables like an oil price index which controls for such input prices.

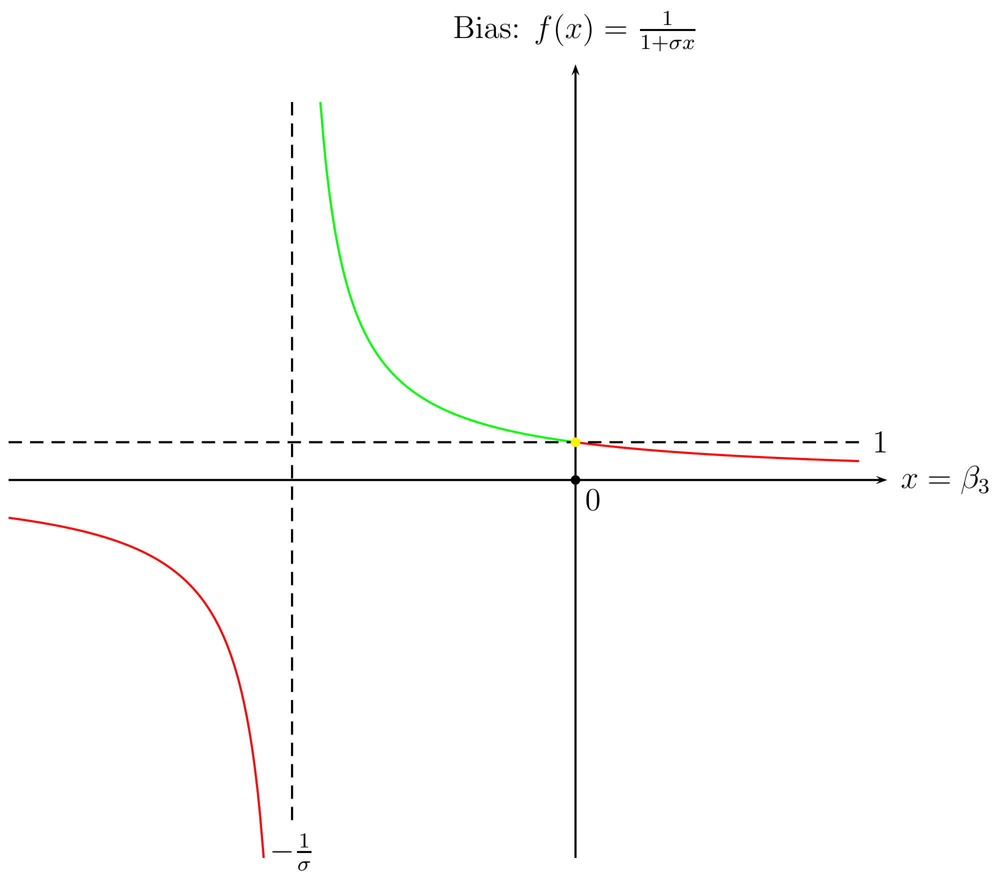
Of course, it remains an empirical question if the propositions of this paper hold. Recent work by Novy (2007) shows how to calculate the mean bilateral trade cost factor from the theoretical gravity equation (8). Jacks, Meissner, and Novy (2008) use this approach to regress (average) trade costs on the usual variables, but they do not control for trade volume, yet. This literature could be a starting point to study the relationship between average trade costs and the trade volume.

A Visualization of the bias term

A.1 Bias term as a rational function



A.2 Bias term with a given value for the elasticity of substitution σ



References

ANDERSON, J. E. (1979): "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 69(1), 106–16.

ANDERSON, J. E., AND E. VAN WINCOOP (2003): "Gravity with Gravitas: A solution to the Border Puzzle," *American Economic Review*, 93(1), 170–92.

——— (2004): "Trade Costs," *Journal of Economic Literature*, 42, 691–751.

ARMINGTON, P. S. (1969): "A Theory of Demand Distinguished by Place of Production," *IMF Staff Papers*, 16(1), 159–178.

BAIER, S. L., AND J. H. BERGSTRAND (2001): "The growth of world trade: tariffs, transport costs, and income similarity," *Journal of International Economics*, 53(1), 1–27.

BERGSTRAND, J. H. (1985): "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence," *Review of Economics and Statistics*, 67(3), 474–81.

——— (1989): "The Generalised Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade," *Review of Economics and Statistics*, 71(1), 143–53.

——— (1990): "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-Industry Trade," *Economic Journal of Economic Literature*, 100, 1216–29.

Brun, J.-F., C. Carrère, P. Guillaumont, and J. de Melo (2005): "Has Distance Died? Evidence from a Panel Gravity Model," *World Bank Economic Review*, 19(1), 99–120.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–1721.

Deardorff, A. V. (1998): "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical Work?," in *The Regionalization of the World Economy*, ed. by J. A. Frankel, pp. 7–22. The University of Chicago Press, Chicago.

Eaton, J., and S. Kortum (2002): "Technology, Geography and Trade," *Econometrica*, 70(5), 1741–1779.

Grossman, G. M. (1998): "Comment on A. V. Deardorff, Determinants of International Trade: Does Gravity Work in a Neoclassical World?," in *The Regionalization of the World Economy*, ed. by J. A. Frankel, pp. 29–31. The University of Chicago Press, Chicago.

Helpman, E., and P. R. Krugman (1985): *Market Structure and International Trade: Increasing Returns, Imperfect Competition and the International Economy*. MIT Press, Cambridge.

Helpman, E., M. Melitz, and Y. Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123(2), 441–487.

Hummels, D. (1999): "Towards a Geography of Trade Costs," Working Paper, Purdue University.

——— (2007): "Transportation Costs and International Trade in the Second Era of Globalization," *Journal of Economic Perspectives*, 21(3), 131–154.

Hummels, D., V. Lugovskyy, and A. Skiba (2009): "The trade reducing effects of market power in international shipping," *Journal of Development Economics*, 89(1), 84–97.

Hummels, D., and A. Skiba (2002): "A Virtous Circle? Regional Tariff Liberalization and Scale Economies in Transport," Working Paper, Purdue University.

Jacks, D. S., C. M. Meissner, and D. Novy (2008): "Trade Costs, 1870–2000," *American Economic Review*, 98(2), 529–534.

Jones, R., and H. Kierzkowski (1990): "The Role of Services in Production and International Trade: A Theoretical Framework," in *The Political Economy of International Trade, Essays in Honor of Robert E. Baldwin*, ed. by R. W. Jones, and A. O. Krueger, pp. 31–48. Blackwell, Oxford.

Krugman, P. R. (1979): "Increasing returns, monopolistic competition, and international trade," *Journal of International Economics*, 9(4), 469–479.

Limao, N., and A. J. Venables (2001): "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade," *World Bank Economic Review*, 15(3), 451–479.

McCallum, J. (1995): "National Borders Matter: Canada-U.S. Regional Trade Patterns," *American Economic Review*, 85(3), 615–623.

MELITZ, M. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

NOVY, D. (2007): "Gravity Redux: Measuring International Trade Costs with Panel Data," Warwick Economic Research Papers No. 861.

OBSTFELD, M., AND K. ROGOFF (2001): "The Six Major Puzzles in International Macroeconomics. Is There a Common Cause?," in *NBER Macroeconomics Annual 2000*, ed. by B. S. Bernanke, and K. Rogoff, pp. 339–390.

PÖYHÖNEN, P. (1963): "A Tentative Model for the Volume of Trade between Countries," *Weltwirtschaftliches Archiv*, 90, 93–99.

REDDING, S., AND A. J. VENABLES (2002): "The Economics of Isolation and Distance," *Nordic Journal of Political Economy*, 28, 93–108.

ROSE, A., AND E. VAN WINCOOP (2001): "National Money as a Barrier to Trade: The Real Case for Currency Union," *American Economic Review*, 91(2), 385–90.

SAMUELSON, P. (1954): "The Transfer Problem and Transport Costs: Analysis of Effects of Trade Impediments," *Economic Journal*, 64, 264–289.

TINBERGEN, J. (1962): *Shaping the World Economy*. Twentieth Century Fund, New York.

Dresden Discussion Paper Series in Economics

04/08 **Bieta, Volker / Broll, Udo / Siebe, Wilfried:** The Banking Firm: The Role of Signaling with Collaterals

05/08 **Felder, Stefan / Werblow, Andreas:** Do the Age Profiles of Health Care Expendidure Really Steepen over Time? New Evidence from Swiss Cantons

06/08 **Broll, Udo / Wahl, Jack E. / Wessel, Christoph:** Export Production, Hedging Exchange Rate Risk: The Duopoly Case

07/08 **Choi, Jay-Pil / Thum, Marcel:** The Economics of Political-Connected Firms

08/08 **Broll, Udo / Gilroy, B. Michael / Lukas, Elmar:** Export Production under Exchange Rate Uncertainty

09/08 **Broll, Udo / Schmidt, Christian W.:** The Effect of Exchange Rate Risk on US Foreign Direct Investment: An Empirical Analysis

10/08 **Lehmann-Waffenschmidt, B. Cornelia:** An Evolutionary Agent-Based Simulation Model for the Industry Lifecycle

11/08 **Mukherjee, Vivekananda / Ramani, Shyama V.:** R&D Cooperation, Asymmetrical Technological Capabilities and Rationale for Technology Parks

12/08 **Lehmann-Waffenschmidt, Marco / Roth, Gerhard / Thießen, Friedrich:** Die (innere) Logik des Entscheidens – Zur neurobiologischen Fundierung ökonomischer Entscheidungen

13/08 **Dopfer, Kurt:** Zur Neubestimmung von Geschichte und Theoriegeschichte in den Wirtschaftswissenschaften – Ko-evolution von Wirtschaftsbeobachtung und Wirtschaftswirklichkeit

14/08 **Lessmann, Christian / Markwardt, Gunther:** One Size fits all? Decentralization, Corruption, and the Monitoring of Bureaucrats

15/08 **Farzanegan, Mohammad Reza / Markwardt, Gunther:** The Effects of Oil Price Shocks on the Iranian Economy

16/08 **Fuchs, Michaela / Wohlrabe, Klaus:** Instituitions, Trade and Integration: What can be expected within the enlarged EU?

01/09 **Broll, Udo / Egozcue, Martin / Wong, Wing-Keung:** Prospect Theory and Two Moment Model: the Firm under Price Uncertainty

02/09 **Broll, Udo / Eckwert, Berhard:** Modelling Information and Hedging: the Exporting Firm

03/09 **Binswanger, Hans Christoph:** Die Wachstumsspirale in der Krise – Ansätze zu einem nachhaltigem Wachstum

04/09 **Brunow, Stefan / Hirte, Georg:** Regional Age Structure and Economic Growth: An Econometric Study for German Regions

05/09 **Broll, Udo / Kemnitz, Alexander / Mukherjee, Vivekananda:** Globalization and a Welfare Program for the Marginalized

06/09 **Tscharaktschiew, Stefan / Hirte, Georg:** An Urban General Equilibrium Model with Multiple Household Structures and Travel Mode Choice

07/09 **Tscharaktschiew, Stefan / Hirte, Georg:** How does the Household Structure Shape the Urban Economy?

08/09 **Lessmann, Christian:** Fiscal Decentralization and Regional Disparity: Evidence from Cross-section and Panel Data

09/09 **Lessmann, Christian / Markwardt, Gunther:** Aid, Growth and Decentralization

10/09 **Broll, Udo / Wahl, Jack E. / Wessel, Christoph:** Export and Benefits of Hedging in Emerging Economies

11/09 **Rudolph, Stephan:** The Gravity Equation with Micro-Founded Trade Costs